



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 14, 2017 – 09:55 PM EDT

PDB ID : 2GWS  
Title : Crystal Structure of human DNA Polymerase lambda with a G/G mismatch in the primer terminus  
Authors : Garcia-Diaz, M.; Picher, A.J.; Bebenek, K.; Pedersen, L.C.; Kunkel, T.A.; Blanco, L.  
Deposited on : unknown  
Resolution : 2.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

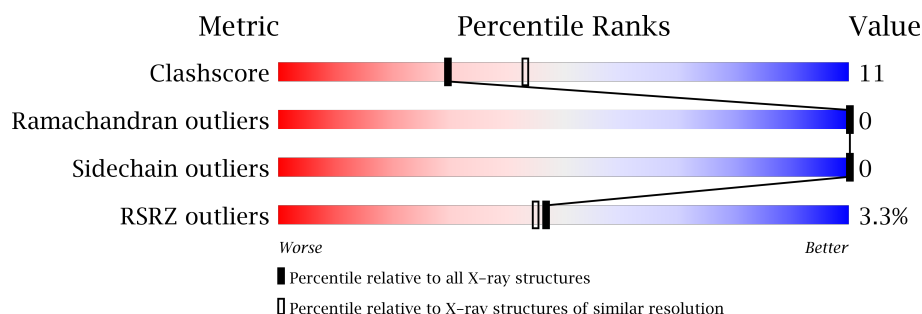
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



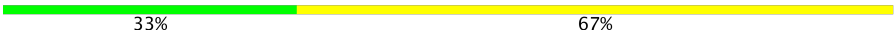

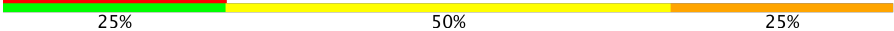

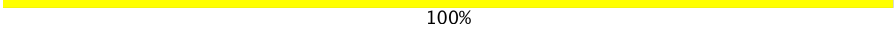




Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	F	11	<div> <div>18%</div> <div>82%</div> </div>
1	J	11	<div> <div>27%</div> <div>73%</div> </div>
1	N	11	<div> <div>18%</div> <div>82%</div> </div>
1	T	11	<div> <div>45%</div> <div>55%</div> </div>
2	G	6	<div> <div>33%</div> <div>67%</div> </div>
2	K	6	<div> <div>50%</div> <div>50%</div> </div>
2	P	6	<div> <div>83%</div> <div>17%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	R	6	 33% 67%
3	D	4	 50% 25% 25%
3	H	4	 25% 25% 50% 25%
3	L	4	 75% 25%
3	Q	4	 100%
4	A	335	 83% 15% .%
4	E	335	 73% 24% 7% .
4	I	335	 72% 25% 5% .
4	M	335	 84% 14% .

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NA	A	940	-	-	-	X
8	EDO	E	2802	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 12600 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a DNA chain called 5'-D(\*CP\*GP\*GP\*CP\*AP\*GP\*CP\*GP\*CP\*AP\*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	11	Total	C	N	O	P	0	0	0
			222	105	45	62	10			
1	F	11	Total	C	N	O	P	0	0	0
			222	105	45	62	10			
1	J	11	Total	C	N	O	P	0	0	0
			222	105	45	62	10			
1	N	11	Total	C	N	O	P	0	0	0
			222	105	45	62	10			

- Molecule 2 is a DNA chain called 5'-D(\*GP\*TP\*GP\*CP\*GP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	6	Total	C	N	O	P	0	0	0
			124	59	25	35	5			
2	G	6	Total	C	N	O	P	0	0	0
			124	59	25	35	5			
2	K	6	Total	C	N	O	P	0	0	0
			124	59	25	35	5			
2	R	6	Total	C	N	O	P	0	0	0
			124	59	25	35	5			

- Molecule 3 is a DNA chain called 5'-D(P\*GP\*CP\*CP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			
3	H	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			
3	L	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			
3	Q	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			

- Molecule 4 is a protein called DNA polymerase lambda.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	327	Total	C	N	O	S	0	2	0
			2562	1607	466	477	12			
4	E	326	Total	C	N	O	S	0	0	0
			2446	1538	441	455	12			
4	I	326	Total	C	N	O	S	0	1	0
			2514	1576	459	467	12			
4	M	326	Total	C	N	O	S	0	0	0
			2530	1588	460	470	12			

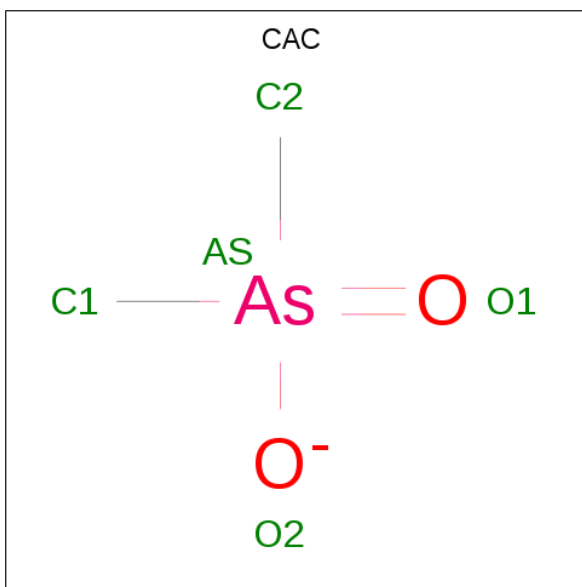
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	241	MET	-	CLONING ARTIFACT	UNP Q9UGP5
E	241	MET	-	CLONING ARTIFACT	UNP Q9UGP5
I	241	MET	-	CLONING ARTIFACT	UNP Q9UGP5
M	241	MET	-	CLONING ARTIFACT	UNP Q9UGP5

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	P	1	Total	Na	0	0
			1	1		
5	Q	1	Total	Na	0	0
			1	1		
5	E	1	Total	Na	0	0
			1	1		
5	I	1	Total	Na	0	0
			1	1		
5	A	1	Total	Na	0	0
			1	1		
5	T	1	Total	Na	0	0
			1	1		
5	M	2	Total	Na	0	0
			2	2		

- Molecule 6 is CACODYLATE ION (three-letter code: CAC) (formula: C<sub>2</sub>H<sub>6</sub>AsO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	D	1	Total	As	C	O	0	0
			5	1	2	2		
6	H	1	Total	As	C	O	0	0
			5	1	2	2		
6	J	1	Total	As	C	O	0	0
			5	1	2	2		

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	F	1	Total	Mg	0	0
			1	1		

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	N	1	Total	C	O	0	0
			4	2	2		
8	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	Cl	0	0
			1	1		

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	T	43	Total	O	0	0
			43	43		
10	P	17	Total	O	0	0
			17	17		
10	D	7	Total	O	0	0
			7	7		
10	F	17	Total	O	0	0
			17	17		
10	G	7	Total	O	0	0
			7	7		
10	H	4	Total	O	0	0
			4	4		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	J	18	Total 18	O 18	0	0
10	K	10	Total 10	O 10	0	0
10	L	4	Total 4	O 4	0	0
10	N	35	Total 35	O 35	0	0
10	R	16	Total 16	O 16	0	0
10	Q	6	Total 6	O 6	0	0
10	A	206	Total 206	O 206	0	0
10	E	77	Total 77	O 77	0	0
10	I	89	Total 89	O 89	0	0
10	M	243	Total 243	O 243	0	0

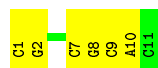


### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

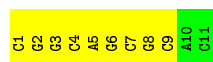
- Molecule 1: 5'-D(\*CP\*GP\*GP\*CP\*AP\*GP\*CP\*GP\*CP\*AP\*C)-3'

Chain T: 



- Molecule 1: 5'-D(\*CP\*GP\*GP\*CP\*AP\*GP\*CP\*GP\*CP\*AP\*C)-3'

Chain F: 



- Molecule 1: 5'-D(\*CP\*GP\*GP\*CP\*AP\*GP\*CP\*GP\*CP\*AP\*C)-3'

Chain J: 




- Molecule 1: 5'-D(\*CP\*GP\*GP\*CP\*AP\*GP\*CP\*GP\*CP\*AP\*C)-3'

Chain N: 



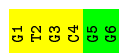
- Molecule 2: 5'-D(\*GP\*TP\*GP\*CP\*GP\*G)-3'

Chain P: 



- Molecule 2: 5'-D(\*GP\*TP\*GP\*CP\*GP\*G)-3'

Chain G: 



- Molecule 2: 5'-D(\*GP\*TP\*GP\*CP\*GP\*G)-3'



- Molecule 2: 5'-D(\*GP\*TP\*GP\*CP\*GP\*G)-3'



- Molecule 3: 5'-D(P\*GP\*CP\*CP\*G)-3'



- Molecule 3: 5'-D(P\*GP\*CP\*CP\*G)-3'



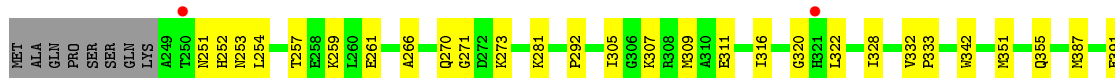
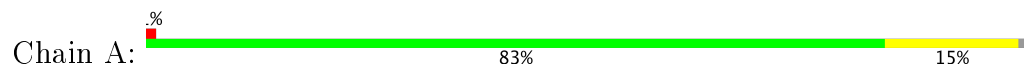
- Molecule 3: 5'-D(P\*GP\*CP\*CP\*G)-3'

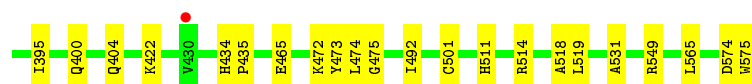


- Molecule 3: 5'-D(P\*GP\*CP\*CP\*G)-3'

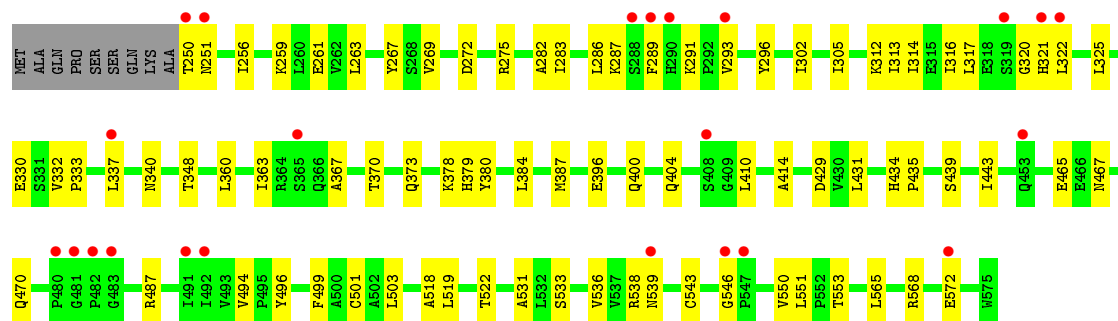
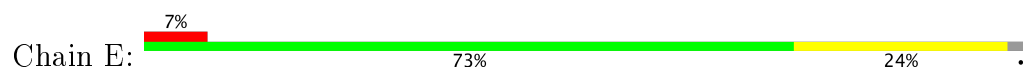


- Molecule 4: DNA polymerase lambda

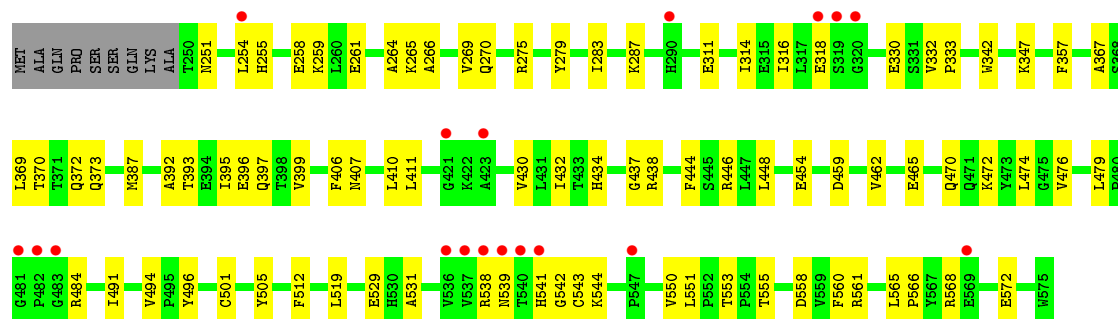




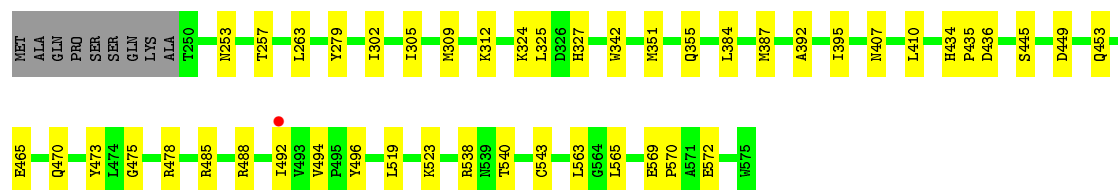
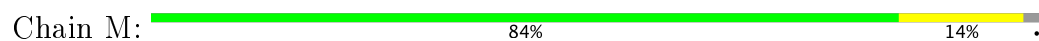
• Molecule 4: DNA polymerase lambda



• Molecule 4: DNA polymerase lambda



• Molecule 4: DNA polymerase lambda



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	192.37Å 98.27Å 104.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.40 28.50 – 2.41	Depositor EDS
% Data completeness (in resolution range)	92.2 (50.00-2.40) 93.0 (28.50-2.41)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.50 (at 2.42Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.209 , 0.253 0.210 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	36.9	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 44.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12600	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.51% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CAC, NA, MG, EDO, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	F	0.32	0/249	0.68	0/382
1	J	0.30	0/249	0.69	0/382
1	N	0.38	0/249	0.77	0/382
1	T	0.45	0/249	0.79	0/382
2	G	0.31	0/139	0.60	0/214
2	K	0.32	0/139	0.63	0/214
2	P	0.40	0/139	0.68	0/214
2	R	0.50	0/139	0.70	0/214
3	D	0.96	1/92 (1.1%)	0.76	0/138
3	H	0.96	1/92 (1.1%)	0.78	0/138
3	L	0.93	1/92 (1.1%)	0.74	0/138
3	Q	0.96	1/92 (1.1%)	0.74	0/138
4	A	0.37	0/2614	0.59	0/3531
4	E	0.34	0/2495	0.54	0/3385
4	I	0.35	1/2566 (0.0%)	0.52	0/3474
4	M	0.38	0/2583	0.62	0/3493
All	All	0.39	5/12178 (0.0%)	0.60	0/16819

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	L	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	1	DG	OP3-P	-7.26	1.52	1.61
3	Q	1	DG	OP3-P	-7.19	1.52	1.61

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	1	DG	OP3-P	-6.93	1.52	1.61
3	H	1	DG	OP3-P	-6.92	1.52	1.61
4	I	265	LYS	CD-CE	5.50	1.65	1.51

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	L	1	DG	Sidechain

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	222	0	123	9	0
1	J	222	0	123	8	0
1	N	222	0	123	7	0
1	T	222	0	123	6	0
2	G	124	0	69	5	0
2	K	124	0	69	4	0
2	P	124	0	69	1	0
2	R	124	0	69	5	0
3	D	83	0	45	1	0
3	H	83	0	45	2	0
3	L	83	0	45	0	0
3	Q	83	0	45	2	0
4	A	2562	0	2527	34	0
4	E	2446	0	2358	67	0
4	I	2514	0	2441	66	0
4	M	2530	0	2484	37	0
5	A	1	0	0	0	0
5	E	1	0	0	0	0
5	I	1	0	0	0	0
5	M	2	0	0	0	0
5	P	1	0	0	0	0
5	Q	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	T	1	0	0	0	0
6	D	5	0	0	0	0
6	H	5	0	0	0	0
6	J	5	0	0	0	0
7	F	1	0	0	0	0
8	E	4	0	6	0	0
8	N	4	0	6	0	0
9	A	1	0	0	0	0
10	A	206	0	0	1	0
10	D	7	0	0	0	0
10	E	77	0	0	0	0
10	F	17	0	0	1	0
10	G	7	0	0	0	0
10	H	4	0	0	0	0
10	I	89	0	0	1	0
10	J	18	0	0	0	0
10	K	10	0	0	0	0
10	L	4	0	0	0	0
10	M	243	0	0	4	0
10	N	35	0	0	0	0
10	P	17	0	0	0	0
10	Q	6	0	0	0	0
10	R	16	0	0	0	0
10	T	43	0	0	1	0
All	All	12600	0	10770	241	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 11.

The worst 5 of 241 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:3:DG:H2''	1:F:4:DC:H5'	1.55	0.87
1:N:9:DC:H5'	4:M:465:GLU:HG3	1.58	0.85
4:I:251:ASN:HD22	4:I:287:LYS:HA	1.45	0.81
4:I:251:ASN:ND2	4:I:287:LYS:HA	1.95	0.81
4:M:519:LEU:HD22	4:M:565:LEU:HD11	1.64	0.80

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	A	327/335 (98%)	313 (96%)	14 (4%)	0	100	100
4	E	324/335 (97%)	314 (97%)	10 (3%)	0	100	100
4	I	325/335 (97%)	312 (96%)	13 (4%)	0	100	100
4	M	324/335 (97%)	316 (98%)	8 (2%)	0	100	100
All	All	1300/1340 (97%)	1255 (96%)	45 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	A	269/281 (96%)	269 (100%)	0	100	100
4	E	247/281 (88%)	247 (100%)	0	100	100
4	I	259/281 (92%)	259 (100%)	0	100	100
4	M	265/281 (94%)	265 (100%)	0	100	100
All	All	1040/1124 (92%)	1040 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:



Mol	Chain	Res	Type
4	E	397	GLN
4	I	297	GLN
4	M	397	GLN
4	E	350	GLN
4	I	486	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 15 ligands modelled in this entry, 10 are monoatomic - leaving 5 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
6	CAC	D	949	-	0,4,4	0.00	-	0,6,6	0.00	-
8	EDO	E	2802	-	3,3,3	0.54	0	2,2,2	0.58	0
6	CAC	H	950	-	0,4,4	0.00	-	0,6,6	0.00	-
6	CAC	J	951	-	0,4,4	0.00	-	0,6,6	0.00	-
8	EDO	N	2801	-	3,3,3	0.74	0	2,2,2	0.70	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	CAC	D	949	-	-	0/0/0/0	0/0/0/0
8	EDO	E	2802	-	-	0/1/1/1	0/0/0/0
6	CAC	H	950	-	-	0/0/0/0	0/0/0/0
6	CAC	J	951	-	-	0/0/0/0	0/0/0/0
8	EDO	N	2801	-	-	0/1/1/1	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	F	11/11 (100%)	0.05	0 100 100	46, 51, 58, 58	0
1	J	11/11 (100%)	-0.05	0 100 100	40, 45, 58, 63	0
1	N	11/11 (100%)	-0.47	0 100 100	22, 26, 43, 47	0
1	T	11/11 (100%)	-0.67	0 100 100	22, 28, 38, 39	0
2	G	6/6 (100%)	-0.07	0 100 100	43, 49, 65, 66	0
2	K	6/6 (100%)	0.21	0 100 100	39, 47, 58, 61	0
2	P	6/6 (100%)	-0.37	0 100 100	21, 26, 39, 44	0
2	R	6/6 (100%)	-0.45	0 100 100	18, 24, 35, 36	0
3	D	4/4 (100%)	-0.87	0 100 100	32, 33, 38, 43	0
3	H	4/4 (100%)	0.37	1 (25%) 1 1	52, 58, 58, 61	0
3	L	4/4 (100%)	-0.44	0 100 100	55, 57, 59, 61	0
3	Q	4/4 (100%)	-0.52	0 100 100	32, 33, 36, 37	0
4	A	327/335 (97%)	-0.18	3 (0%) 84 82	17, 32, 62, 76	0
4	E	326/335 (97%)	0.38	23 (7%) 17 15	36, 57, 84, 99	0
4	I	326/335 (97%)	0.29	18 (5%) 26 24	37, 54, 80, 91	0
4	M	326/335 (97%)	-0.28	1 (0%) 93 93	16, 29, 55, 72	0
All	All	1389/1424 (97%)	0.03	46 (3%) 47 45	16, 44, 75, 99	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	E	321	HIS	4.0
4	E	483	GLY	3.8
4	I	320	GLY	3.7
4	I	254	LEU	3.5
4	I	423	ALA	3.5

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
8	EDO	E	2802	4/4	0.72	0.42	7.43	64,66,67,67	0
5	NA	A	940	1/1	0.97	0.19	3.29	25,25,25,25	0
9	CL	A	952	1/1	0.98	0.16	1.53	41,41,41,41	0
6	CAC	J	951	5/5	0.80	0.20	1.33	171,171,172,172	0
5	NA	M	943	1/1	0.95	0.13	0.40	25,25,25,25	0
5	NA	E	941	1/1	0.93	0.10	-1.77	44,44,44,44	0
5	NA	I	942	1/1	0.93	0.08	-2.06	41,41,41,41	0
8	EDO	N	2801	4/4	0.69	0.51	-	38,42,42,44	0
5	NA	Q	948	1/1	0.80	0.17	-	53,53,53,53	0
6	CAC	H	950	5/5	0.90	0.19	-	160,160,160,160	0
5	NA	M	947	1/1	0.94	0.41	-	43,43,43,43	0
5	NA	P	945	1/1	0.98	0.07	-	39,39,39,39	0
6	CAC	D	949	5/5	0.90	0.17	-	142,143,143,143	0
5	NA	T	944	1/1	0.88	0.05	-	43,43,43,43	0
7	MG	F	946	1/1	0.93	0.15	-	56,56,56,56	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.